HELSINGIN YLIOPISTO

LISTENER ACTION IN CONVERSATION

THE CASE OF VERBAL BACKCHANNELLING AND CO-OCCURRING GESTURES

Sonja Korkiakoski
Pro Gradu Thesis
English Philology
Department of Modern Languages
University of Helsinki
April 2017

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Humanistinen tiedekunta	Laitos – Institution – Department Nykykielten laitos	
Tekijä – Författare – Author Sonja Korkiakoski		
Työn nimi – Arbetets titel – Title Listener Action in Conversation – The Case of Verbal Backchannelling and Co-occurring Gestures		
Oppiaine – Läroämne – Subject Englantilainen filologia		
Työn laji – Arbetets art – Level Pro Gradu -tutkielma	Aika – Datum – Month and year Huhtikuu 2017	Sivumäärä– Sidoantal – Number of pages 51 + lähteet ja liitteet

Tiivistelmä – Referat – Abstract

Tutkielma käsittelee kuuntelijan käytöstä keskustelussa: hänen kielellistä sekä eleellistä palautettaan (*backchannelling*). Aineistonani on kolme videoitua keskustelua Scottish Corpus of Text and Speech – korpuksesta (SCOTS), ja niistä samassa korpuksessa olevat litteraatit. Keskustelut ovat noin puolen tunnin mittaisia kahden naisopiskelijan keskusteluja, ja ne sijoittuvat huoneeseen, jossa heidän lisäkseen on kaksi nauhoittajaa. Tarkoituksena on tutkia kolmea asiaa: kulkevatko jotkin suulliset palaute-elementit käsi kädessä tiettyjen eleiden kanssa, vaikuttavatko eleet kyseisten elementtien merkitykseen, ja millaisia kuuntelijoita keskustelijat ylipäätään ovat.

Käytän termiä *palaute* tässä tutkielmassa viittaamaan ilmiöön, jolla kuuntelija verbaalisesti viestii puhujalle, keskeyttämättä hänen puheenvuoroaan, miten keskustelu sujuu. Tällaisen palautteen tärkein tehtävä on rohkaista puhujaa jatkamaan, mutta samalla se voi ilmaista muun muassa kuuntelijan mielenkiintoa kyseiseen puheenaiheeseen. Koska työskentelen multimodaalisen materiaalin kanssa (litteraatti, ääni ja videokuva), käsittelen myös näiden eri materiaalimuotojen tarjoamia mahdollisuuksia aiheeni tutkimuksessa.

Yleisin palaute-elementti aineistossani on *yeah*, joka suurimmaksi osaksi esiintyy samanaikaisesti nyökkäyksen kanssa. Toisiksi suosituin on nauru, jota käyttää neljä kuudesta kuuntelijasta. Naurua ei ole lähdeteoksissani mainittu, joten tämä tulos johtuu joko uudistamastani palautteen määritelmästä, tai siitä, että materiaalinani on 2000-luvulla tuotettu, luonteva nuorten opiskelijoiden puhe. Eleillä kuuntelijat voivat sekä antaa että vahvistaa suullisten palautteen merkitystä: merkitykseltään neutraali *mm* voi esimerkiksi tarkoittaa samaa mieltä olemista yhdistämällä nyökkäys kyseiseen suulliseen elementtiin.

Tutkielma keskittyy lingvistisiin elementteihin ja niiden kanssa samaan aikaan tehtyyn eleelliseen viestintään, joten verbaalisen palautteen prosodisuus jää vähemmälle huomiolle.

Avainsanat – Nyckelord – Keywords
backchannel, backchannelling, listener, listener action, gesture, discourse analysis, conversation, SCOTS, kuuntelija, keskustelu, eletutkimus, diskurssianalyysi

Säilytyspaikka – Förvaringställe – Where deposited

Muita tietoja – Övriga uppgifter – Additional information

TABLE OF CONTENTS

1. INTRODUCTION

1.1. MOTIVATION FOR STUDY

"... [CONVERSATIONALISTS] ARE EXPECTED TO OBSERVE THE BASIC RULES FOR TURNTAKING AND TO 'LISTEN ACTIVELY' WHEN THEY ARE NOT SPEAKING."

(STENSTRÖM 1994, 18)

I started studying English as a foreign language from third grade onwards. We always focused on gaining competence as a speaker and writer, going meticulously over grammar and vocabulary of each new chapter. I learned how to communicate my wishes in everyday situations. In upper secondary school the vocabulary got more complex and I found that I could express myself in a more accurate manner. However, I had not been introduced to the *listener action* phenomenon until I started to study for University entrance exams. In one of the required books Plag et al. define *backchannels* as "a signal to the speaker that the hearer is still there and listening" (2007, 172). The idea of listener importance was quite soon forgotten, only to resurface again when it came time to think about a topic for this Pro Gradu thesis.

But how important can the listener's actions really be? Showing interest towards the speaker and their topic is polite, of course, but the number of backchannelling instances have been found to have an effect on conversations, namely in the length of the speaker's turn. For instance, Pam Peters and Deanna Wong claim that there is a "strong positive correlation between the number of backchannels received and the length of a speaker's turn" (2015, 417), and the magnitude of this effect can be seen in Christoph Rühlemann's study on backchannels in storytelling. He found that in storytelling situations recorded in the British National Corpus, if no backchannelling occurs, the speaker's turn tends to be half the length compared to those turns that receive backchannelling (2016). In the light of these results backchannel studies can provide important information about discourse and

affecting interpersonal relationships – since we do feel more comfortable talking with people who listen and show it.

In a constantly globalizing world it is important to be able to communicate with people from all kinds of backgrounds. While English, as a *lingua franca* (e.g. Seidlhofer 2005, 339), is increasingly one of the most globally used language, we all use it in different ways. Leaving out the obvious differences of speaker accents and using grammar structures that are sometimes heavily influenced by the mother tongue, there are some differences in *listener actions* that can be found as well.

> "THE POINT IS THAT SWEDES GRUNT DIFFERENTLY (AND POSSIBLY LESS) THAN SPEAKERS OF ENGLISH. THUS IT IS A COMMON EXPERIENCE FOR A SWEDE TO SIGNAL ACQUIESCENCE OR AGREEMENT BY PRONOUNCING [MM] WITH RISING INTONATION AND THEN HAVE AN ENGLISH-SPEAKING INTERLOCUTOR REPEAT HIS/HER PREVIOUS UTTERANCE, THINKING THAT THE SWEDE HAD SIGNALED NON-COMPREHENSION"
>
> (TOTTIE 1989, 270)

Backchannels are the listener's way of managing discourse, most commonly perhaps to facilitate it. Differences in cultural and linguistic conventions, however, can cause misunderstandings such as described in the quote above from Gunnel Tottie. Incidentally those differences cannot be eliminated, since there are approximately 2 billion speakers of English (Crystal 2008, 5), out of which around 339 million are native speakers (Ethnologue website). On top of the differences between native and non-native speakers, there can be divergence in the language use between different varieties of English (e.g. Wong & Peters 2007). John Gumperz concurs, saying "[...] sociolinguistic research [...] has demonstrated [...] that all existing human communities are diverse at all levels of linguistic structure [...]" (1982, 324).

In conclusion, when studying something that draws from the contextual pool of linguistic and cultural background of the listener, in this case *backchannels*, a good starting point is to focus on the instances that appear in the data itself instead of focusing on a specific ready-made list that has arisen from other studies using other data. It is important to be aware that backchannels have not been strictly defined, and the researcher is responsible for clearly stating how they define, use, and research them. Backchannels are context bound and variable, and as this thesis shows, also *multimodal*.

1.2. RESEARCH QUESTIONS

The aim of this thesis is to refine the elusive definition of the phenomenon called *backchannel*, look for instances of them in my dataset, identify the co-occurring gestures and make conclusions about my data based on my findings. Thus, my research questions are:

Q1. Do some verbal backchannels co-occur with certain gestures in my data? Are there patterns of co-occurrence that would predict these combinations?

Q2. Do the verbal backchannels and gestures enhance or give each other meaning, and if yes, how?

Q3. What does backchannel use say about the listeners?

Reasoning for question one, searching for patterns of co-occurrence of verbal backchannels and gestures, is that it gives a more detailed picture of listener actions and their influence. If we only focus on transcriptions of a conversation, without having access to video or audio, we might miss out on nuances that in the conversational context 'make or break' the mood, not to mention gestures that "[...] *complement[s]* the expression achieved in words." (Kendon 2004, 116). These are considered in question two. For example, let us take an extremely common and much researched backchannel *mm* (e.g. in Peters and Wong 2015, Gardner 1998, Kjellmer 2009). Since it does not carry meaning in itself, its intention is up for interpretation, and it is here gestures can make a world of difference. Whereas a nod, a common gesture of agreement, would give *mm* a similar meaning (or interpretation) to *yeah*, glancing away creates emotive distance and thus indicates disagreement or disinterest. I look at these kinds of combinations in my material in both Analysis and in Discussion and Conclusions.

What comes to the patterns of co-occurrence I compile a table for each informant based on their backchannelling in the Analysis section. These tables include a three of their most used backchannels and their co-occurring gestures, and

I also explore their overall listener action and participation in the interaction. Therefore, the third research question is explored in the Analysis.

2. THEORY

In the first section I argue for my decision to use the term 'backchannel' in this thesis by comparing it with its relative concepts of *minimal response* and *receipt token*. Then I define the term how I use it in this thesis. In the second section I give an overview of my multiperspectival approach, including a short discussion on Discourse Analysis and one of its branches, Discursive Psychology. Lastly, the Theory section is concluded with a few remarks about Gesture Studies.

2.1. BACKCHANNELS

The nomenclature of studies concerning verbal listener action is varied and overlapping. There are some terms, however, that keep popping up again and again, and one of them is *minimal response*. In 1986 Jennifer Coates uses it in her book *Women, Men and Language* to describe a short listener response. Later she acknowledges the listener's active role in a conversation, and that these responses can be used to not only support the speaker, but also to signal a lack of understanding or interest (100-101). It is important to acknowledge this fact because listener action affects discourse, and disregarding the negative effects leaves the study partial and overly highlights the listener's role as an encourager rather than a regulator. Due to the nature of the data used in this thesis (discussed in Material and Methods), however, I come across mainly continuer backchannels than regulating ones.

Along the way there have been multiple definitions, mostly short word lists, of *minimal responses*. For example, Don Zimmerman and Candace West suggest that responses such as *uh huh*, *yeah*, and *um hmm* (in later studies replaced by the orthographical form *mhmm)*, monitor the speaker, and that they can be accompanied by nonverbal cues, such as

a nod (1975, 108). Suggesting that the listener can monitor the speaker may at first sound absurd since the speaker is seen having the power in the conversation, but as has already been mentioned above, listener action can affect and for example the length of the speaker's story (Rühlemann 2016).

Interestingly enough, studies of *minimal responses* have focused on gender differences (e.g. Coates 1986, Coates 2016 and Zimmerman & West 1975) while the focus of *backchannel* studies seems to be on their influence on discourse, and the differences between Englishes (e.g. Rühlemann 2016 and Wong & Peters 2007, respectively). Research has shown that women use *minimal response*s more than men, and that while they do the 'interactional shitwork', as Coates quotes Fisherman, men use same responses in a different manner to reinforce male dominance (2016, 87-88). It certainly seems that studying women's backchannelling (at least the *mm*s and *uh-huh*s) is more lucrative, but some of these studies have been conducted some time ago.

In 2014 Coates suggests that the term *backchannel*, which has gained popularity among researchers, could be used as an alternative to *minimal response*, although limiting the definition to a list of minimal utterances "such as *yeah* or *right* or *mhm*" (87). I find the term *minimal response* limiting, since listener action can, and *is*, much more than a few hums in-between speaker utterances. In my material there are a couple of long repetitions of the speaker's utterance, spoken in disbelief, or a comment to the speaker's topic, almost, but not quite, a speaking turn in itself. And while these minimal responses can and do have backchannelling functions, they should not be the only defining factor.

Receipt token is another term that pops up quite often in similar studies, but is not explicitly explained (e.g. Gardner 1998). The name, however, strongly implies an automatic response ('receipt') without meaning or substance in itself ('token'), although it is often used as an interchangeable term with *minimal response* and *backchannel*. Although backchannelling comes naturally and for example I only realize I am doing it while thinking about this thesis, it does not mean it cannot be done consciously to guide the discourse for certain purposes. Backchannelling in interviews, for example, aims to keep the guest talking. This is why the term meaning vocal listener action should not innately refer to something automatic or not having any meaning or relevance – and thus I do not use *receipt token* when describing backchannelling items. I use the term *backchannel item* here to refer to the word of a part of an utterance that functions as a backchannel.

Victor Yngve is credited to have first introduced the term *back channel* in the sense of verbal listener action, during the Sixth Regional Meeting of the Chicago Linguistic Society in 1970 (see Sorjonen 2001, Kjellmer 2009, Gardner 1998). Many people have built upon this term: for example, Tottie distinguishes between different backchannel functions, such as supportive and regulative, and Stenström ranges backchannels on a scale from indifference to strong involvement (Kjellmer 2009, 84). This idea of involvement shows how excited or interested the listener is which in turn influences the speaker to either continue or perhaps condense their turn so as not to bore the listener.

There has been research on listener action and behaviour before Yngve's paper *Getting a word in edgewise*, however, and some have even tapped into the backchannelling phenomenon. For example, in 1968, Allen T. Dittmann and Lynn G. Llewellyn studied listener responses in the forms of head nods and vocalizations. They found that 51% of single responses, that is, when only a nod or a vocalization is present, were signals of attention rather than serving interpersonal functions such as questioning or responding to speaker's signal for feedback. I deduct from their description that at least those 51% were backchannels, and possibly some feedback signals as well. Dittmann and Llewellyn also found that 30% of joint responses (where both a nod and vocalization were present) were signals of continuing attention – in other words, backchannels. These results, although conducted in a laboratory setting in 1968 and the respondents consisted of only 20 college students, would indicate that co-occurrences of verbal and gestural backchannelling is of interest and worth studying.

There are lists that aim to define backchannels. These might be useful up to a certain point, e.g. when only talking about minimal responses, studying copious amount of data, or when only getting to know what the phenomenon is about. Bjørge introduces a list that includes some examples from different levels of backchannelling:

> "(A) NON-VERBAL MANIFESTATIONS: HEAD NODS. (B) NON-LEXICAL ITEMS: MHM, AH, OH. (C) LEXICAL, PHRASAL AND SYNTACTIC ITEMS: ABSOLUTELY, BRILLIANT, CERTAINLY, DEFINITELY, EXACTLY, EXCELLENT, FINE, GOOD, GREAT, I SEE, OF COURSE, OK, PERFECT, QUITE, REALLY, RIGHT, SO, SURE, THAT'S NICE/RIGHT/NOT BAD, TRUE, YES/ YEAH, YES I KNOW. (D) REPETITION OF OTHER SPEAKER'S UTTERANCE."
>
> (BJØRGE 2009, 196)

The first level or category is the gestural side of, or nonverbal as opposed to verbal, backchannelling. Non-lexical and some lexical items, such as *yeah*, fit into a *minimal response* category. Repetition of the speaker's utterance is included here and is, along with phrasal and syntactic items, what differentiates backchannels Coates' *minimal responses*. *Yes, I know* is an example of a backchannel that looks like a turn, but is here to most likely indicate that answer-like backchannels to *you know?* questions are not turns. The question here is not a typical one (one that requires an answer), but rather a request for reassurance that the listener understands and is on board with the discussion.

One item that Bjørge explicitly leaves out in this list is the word *no* and I completely disagree with that decision, although her reasoning is understandable: "However, as no is used to object to a statement it impacts on the following turn rather than functioning as a turn-continuer, and frequently represents an unsuccessful bid for the floor" (2009, 196). It is true that *no* can indeed denote an objection, but it also is a very good way of asserting empathy, agreement with the speaker (but disagreeing with the topic) and involvement with the speaker and their topic. This, in turn, encourages the speaker to continue, which is one function of backchannels – a *regulative* one (Pipek 2007, 20). It is important to do close reading on the material before analysing it, and same goes for possible backchannel items. For example *mm*, a very common example of a backchannel, is sometimes just an answer. This is why ruling out *no* as a backchannel is hasty. I concur with Tottie's statement that "[...] backchannel status can be determined only on the basis of the following utterance" (Pipek 2007, 21 quotes Tottie 1991, 260).

Instead of calling a list of items backchannels it would be more fruitful to begin from some features and the *functions* of listener action – what do these backchannel items actually *do*? Concerning features, Wong and Peters state that if nothing else, a large group of researchers agree on at least two points. First, *backchannels* occur while the speaker is speaking. Some broaden the definition of speech to speaker's pauses, so non-interruptive listener backchannelling can occur during speaker pauses as well. Second, it is commonly agreed upon that backchannelling does not interrupt the speaker's turn (Wong and Peters 2007, 485-6). Vojtěch Pipek has summarised three features often attributed to *backchannels*: they are not turns, they do not bring any new information, and usually they overlap with the speaker's utterance (2007, 44). However, if one is to view backchannels as listener action that intends to regulate interaction, one important point is to be made:

"MOREOVER, WHAT STARTS AS A BACKCHANNEL MAY END UP AS A TURN, IF THE PREVIOUS SPEAKER SHOWS NO WILLINGNESS TO CONTINUE SPEAKING."

(PIPEK 2007, 20, QUOTING TOTTIE 1991, 257)

Backchannel features, then, can and do change depending on different contexts and on who you ask. Here I concur with the claims that backchannels are not turns, and although they may comment on a topic, they do not bring any relevant new information. The most important thing to consider, however, is that they are defined by the following turn. Just as the listener has the power to guide the conversation, the speaker can go along with it or choose not to.

Backchannels have different functions. The most recognizable perhaps is the *continuer* function (e.g. Piper 2007, Kjellmer 2009, Gardner 1998, Peters & Wong 2015), which encourages the speaker to continue, and sometimes "[...] to provide feedback on how the message is being received" (Knight 2011, 90). Backchannels are also used to show *agreement* (also called consonance or convergence tokens, see Piper 2007, Knight 2011) in forms such as *yeah,* and *interest* or *involvement*, also in the form of the aforementioned *no* to show, for example, surprise or anger**.** Something to be aware of is that backchannels can have different *degrees* of involvement, agreement and interest: a quiet *mm* while looking at something else than the speaker conveys a low degree of involvement and interest, and maybe even disagreement. It is important to note that although I do not look closely at tone, pitch or vowel length, these add meaning and function to backchannels as well.

Seeing as many of backchannel functions overlap I do not especially look for or mark the exact functions of the backchannels I come across in my research. However, I will be looking for items or utterances that function as continuers, feedback about how the information has been received, agreement, acknowledgement of the speaker or their turn, and show of interest and/or involvement. Repetition of the speaker's utterance, or parts of it, is, depending on the situation, either feedback for disrupted or surprising information, or a show of interest or involvement. I do not consider 'missing' backchannels, or where the purposeful lack of backchannelling is used and causes the speaker to stop their turn.

I do not consider 'failed' backchannels in this paper. A backchannel fails if it is followed by the listener taking the turn or if the speaker changes the topic completely. The latter decision is based on the core function of *continuer*, even if it conveys listener noninterest. As previously stated, the reaction to a backchannel is what defines it: if the

speaker keeps going on about a topic the listener has expressed noninterest in, the backchannel has functioned as a continuer. What comes to agreement, answers (even short ones) for straight questions are not included. This is why I do not automatically count every instance of *mmhm* or *yeah* as backchannels: they might just as well be answers, and as such, turns. An answer is contributing to the topic itself, whereas backchannels are a continuer and a feedback function about the *discourse* overall.

Rühlemann describes backchannels as a *discourse phenomenon* (Kjellmer 2009, 83) and I concur with this claim. By using the term 'backchannel' to denote a *phenomenon* or function rather than a list of responses or utterances we avoid the pitfall of missing out on possibly interesting and relevant instances, such as a complex utterance acting as a signal of agreement.

"THESE ISSUES INDICATE THAT THERE IS A VERY REAL NEED TO RETURN TO THE FUNDAMENTALS OF IDENTIFICATION AND DESCRIPTION [...]"

(WONG AND PETERS 2007, 480)

In conclusion, in this paper I define a backchannel as a verbal listener action that functions as a continuer, acknowledges the speaker and/or gives feedback or agreement, and shows a varying degree of interest or involvement (from disinterest to great interest). A backchannel is followed by a speaker continuing their turn, as its function as a continuer suggests and it does not change the topic of conversation. A backchannel is not a straight question as this is overtly leading the conversation and I consider it a turn, but a backchannel can be a tag question: "[t]hey are not really questions but are a way of asking the other person to make a comment and so keep the conversation open" (Brown, English Grammar Secrets webpage). Helping the speaker continue their turn, for example by correcting their incorrect wording or helping them find the correct word is also considered a backchannel here – in the case that the speaker continues, of course.

2.2. MULTIPERSPECTIVAL APPROACH

2.2.1. DISCOURSE ANALYSIS AND DISCURSIVE PSYCHOLOGY

Discourse analysis is concerned with talk-in-interaction, but whereas critical discourse analysis (or CDA) strives for societal change (Jørgensen and Phillips 2002, 64), discursive psychology, which I use in this paper, focuses on specific instances of discourse (7), that is, language use (Schiffrin, Tannen & Hamilton 2001, 1). Both branches have their strengths and limitations. CDA highlights the need for change on the large scale, and therefore there is a need and requirement to generalize the results of research to whole populations in order to suggest how to make a change. Discursive psychology, on the other hand, starts on the grass root level of discourse itself and aims to describe what is happening there. It encourages the researcher to mix different theories and methods, creating a multiperspectival approach – keeping the nature of the data in mind.

When talking about discursive psychology, Jørgensen and Phillips (2002) often refer to Potter and Wetherell, where they "[...] hope to indicate how a new style of socio-psychological research can be erected on the foundations of speech act theory, ethnomethodology and semiotics" (Potter and Wetherell 1987, 32). The strong points of each have a great influence on my thesis and are quite similar in how I perceive backchannels to be. On this I expand next.

In John Austin's Speech Act theory there are three suppositions: that all utterances have meaning, a force and consequences (Potter & Wetherell 1987, 17). The theory has been criticized in that utterances cannot in themselves be categorized to perform certain actions, but that it is in the *response* where we can see what it does (29-30). This problem is avoided by the very nature of my study: as discussed above, the meaning of backchannels is in how they function and how they are interpreted by the speaker – and in this paper, by the researcher. The *form* of a backchannel affects its force: for example *yeah* indicates stronger agreement than *mm* due to its relation to the word *yes*. The *consequences* of listener action can be as intended – but they can also be quite the opposite, and this depends on the speaker's interpretation: for instance in Tottie's example in 1.1 Motivation for Study, a Swede's grunt can be interpreted as a request for clarification by an English speaker.

Since discursive psychology rose from criticism towards realistic and positivist approach, it in itself recognizes the effect of the researcher and the context on the data and

results. It also highlights the importance of leaving the laboratory and observing actual, non-elicited conversation and communication: that is, talk in its natural environment. Ethnomethodology studies just this: how language is used in everyday contexts (Potter & Wetherell 1987, 19). Here the researcher must be especially aware of their assumptions and interpretations and, most importantly, make them *visible*, since language is context-dependent, and we all carry different sets of interpretative tools due to different life experiences and background knowledge. It is imperative the researcher is explicitly aware, and shares this with the audience, that their interpretation affects the results.

Semiotics, on the other hand, asks what is *not* there (Potter & Wetherell 1987, 28). A backchannel can be just as influential on the speaker and their turn as the *lack* of one, as can be seen in Bjørge's study on negotiation (2009). Backchannels are used to direct and regulate discourse: most commonly to facilitate it. In this paper I do not analyse how the lack of a backchannel, for example in the cases of long pauses, affects discourse, but I acknowledge that a backchannel and the lack of one can have an equally strong force.

There are many theories that could be examined here. For instance, a deeper look into Discourse Analysis would show how interactions are structured and how they can be studied, while cooperative principle and recipient design could explain why backchannels are used. Delving deeper into Discourse Analysis would give a more inclusive explanation about its history and evolvement, but since the methods associated with it change according to the purposes of each study and even then they are not exact but rely on researcher interpretation, I do not inspect it closer. Recipient Design is fascinating in itself, but slightly off topic in that while I look at listener action, I do not consider the backchannel effect on speaker and their turn – in this way, the recipient (who in this study is the actually the speaker), is not the subject of this study. Due to the limitations of time and space I do not discuss all the different theories that could link to the backchannelling phenomenon. In this paper I only focus on the backchannels that are used in my data.

2.2.2. GESTURE STUDIES

Anne Kari Bjørge found in her study on English for Special Purposes and the use of backchannelling in ELF negotiations (2009) that in her data 70% of backchannel instances were nods and only 30% were verbal utterances. As seen in section 2.1 Backchannels, her

definition of a backchannel differs somewhat from mine but this does not discount the overflowing evidence of the importance of gestures in her study. Nods are one of the most visible or noted gestures in everyday interactions, but they are not the only ones. At this point it is important to distinguish a gesture from other types of movement. Adam Kendon defines a gesture as "[...] visible action when it is used as an utterance or as a part of an utterance" (2004, 7). Here utterance is understood as an activity that is treated by others as "[...] a communicative 'move', 'turn' or 'contribution'" (7). In his book Kendon explores gestures when they are used as "complements, supplements, substitutes or alternatives to [spoken expressions]" (1). By this definition, gestures themselves can be backchannels, since they can function as utterances, as in the case of a nod meaning agreement. In this paper I use the term *backchannel* exclusively about the verbal instances of the backchannelling phenomenon because they are my starting point and gestures are only looked at in relation to them but I recognize and acknowledge (and even suggest that others do the same) that gestures can and do function as backchannels as well.

Gestures have been of interest from Antiquities to present day, but the amount of attention the field has received in any given time has varied. Quintilianus, for example, wrote on Roman rhetorical doctrine in the first century AD (Kendon 2004, 17). He mainly focused on hands, and his gesture principles made it possible for them to be taught (18-21). In Europe, Bonifacio wrote one of the earliest books dedicated to gestures in 1616 (23).

What is interesting is that while during Antiquity rhetoric and gestures appropriate for that purpose were kept separate from art and theatre, later on these two would learn from each other (Kendon 2004, 32): life imitates art, and art imitates life. Gesture studies, then, can be a valuable source of information on how to imitate emotion on stage. Behaviour is not just affected by emotion, however: for example Gerard de Lairesse, a painting teacher, observed how a person's *habitus* (or education, social class and background) affects behaviour (30). Just as one's educational background affects speaking style, it affects their movements and gestures – their way of being - as well.

Linguists and cultural anthropologists Franz Boas and Edward Sapir studied American Indian languages intending to transform them into written form. In trying to separate behaviour that was and was not significant for a language system, they concluded that a "[...] gesture should be considered part of a broad patterning of communicative behaviour [..]" (Kendon 2004, 65-6). It wasn't, however, until late 1940's when the idea of studying nonverbal communication caught on and even then it took some time before

gestures became the focus of study using audio-visual recording (68-69). This new technology allowed the interrelation of speech and action to be studied (2).

Gestures can be used to get different meanings across. For example, some gestures can be substitutes for utterances: a nod for 'yes', raised middle finger for 'fuck you'. The more insulting the gesture, the more covert variants it has, for example to avoid confrontation. In Germany a gesture meaning 'they are crazy', "forefinger touches the side of the head and is rotated back and forth", is seen as a grave insult, and has resulted in a birth of a gesture that has the same meaning but that can also be mistaken for scratching one's cheek (Kendon 2004, 9).

According to Kendon, nervous habits are not usually accounted for in gesture studies, although they "may sometimes be read by others as symptoms of the individual's moods or feelings" (2004, 8). At first this may seem odd, because the gestures worth studying are those that 'give information', nervousness should absolutely be taken into account since it does give information about the gesturer. The point here, however, seems to be *intention*. Non-intentional or involuntary gestures, such as shivering when it's cold, might not contribute to the talk or conversation itself, but to the surrounding *context*. It is up to other participants to either acknowledge or ignore them. Such gestures can, of course, be used intentionally to signal to others to either close a window or to borrow a jacket. The reasoning behind ignoring involuntary gestures in studies can be found in the context. A person with Tourette's Syndrome, for example, is not taken to mean every twitch of their head as an informational gesture – then why should a person's nervous habit of fiddling with their ring or jewellery? It does not signal a contribution to the conversation itself. Thus, the *gesturalness* of a movement is defined as such by their *context*, just like with backchannels. But as we cannot get inside a speaker's head to see how they interpret listener gestures, we, as observers, are left to our own devices in deciphering whether a gesture is meaningful to them or not.

Seeing as movement and gestures can get different meanings in different contexts, Kendon suggests that "[t]he particular classification systems developed are useful working instruments for a given investigation, but they should not be thought of as more than this" (2004, 85). Few examples of classification systems that work for a given task would be those of Quintilian and Engel's. The former discussed movements minding mainly orators whereas the latter wrote comprehensively about aspects of bodily expression, most useful to actors (Kendon 2004, 85-7). I do not use any ready-made classification systems because the

classification of gestures derives from the data itself - through the subjective eye of the researcher, of course. However, to lend Goffman's concept of *attentional tracks* is useful here. He claims that "an action that is gestural has an immediate appearance of gesturalness" (Kendon 2004, 15), and Kendon cites a study where participants were shown a film with no sound, and were later asked how the character in the film moved or gestured. The first gestures people remembered were seen as deliberate and that seemed to have intention to say something, then movements that sustained or changed bodily position, and after those came the manipulation of objects and nervous or incidental actions. The last two were not seen as communicative, and some used movements in the last attentional track as evidence of the speaker character's mood and were seen as nonintentional (10-15). The results of this study suggest that the more 'gestural' or meaning-making the movement, the easier it is to notice and remember.

Although I look at all visible gestures to find clues about the flow of the conversation, I will only count the visibly meaning-wise relevant ones: for example a nod co-occurring with *yes*, instead of the change of sitting position or a slight twitch of a hand. In some cases there are two or more meaning-making gestures (for example a nod and leaning towards the speaker indicating interest) but from these I choose the one I feel is most relevant in meaning-making or the most visible one - the one I, the observer, pay attention first to or the most.

3. MATERIAL AND METHODS

3.1. SCOTTISH CORPUS OF TEXT AND SPEECH

In this thesis I look at three conversations that occur between two different female students. These data are from the multimodal corpus called the Scottish Corpus of Text and Speech (hereon, SCOTS), and includes a video as well as a transcription of the interaction. SCOTS is a free online resource that consists of over 1300 texts and 4.5 million words. The aim of the

corpus is to represent actual language use, spanning over time from 1945 onwards. It is accessible on http://www.scottishcorpus.ac.uk/.

The project was done in two phases. The first phase was carried out between 2002 and 2004, and was led by teams from University of Glasgow and University of Edinburgh. The former collected text and speech while the latter examined research issues concerning multimodal corpora. The first phase was funded by EPSRC, the Engineering and Physical Sciences Research Council. In a summary of final report submitted to the funder it is mentioned that the collected metadata is more extensive than in many other corpora, and I concur with this claim. The background information about the participants is very extensive, including their level of education, country of birth and where they live, all the way to their parents' careers. Most of this information is not of interest to me in this paper, but for a sociolinguistic study, for example, this kind of metadata is a stroke of luck.

The second phase was carried out by English Language and STELLA project at University of Glasgow. It went on between the years 2004-2007 and was funded by the Arts and Humanities Research Council. Summary of the final report reveals that Google Maps was integrated to the search function, and that the spoken documents have synchronized orthographic transcription. This enables the viewer, for example in the case of video files, to see the video and follow the transcription at the same time.

Most of the publications about the corpus that is written by members of the collecting teams talk of the practicalities of collecting and using a corpus for research, use of metaphors by the informants, and variations of Englishes. The last update was conducted in November 2013, and since the funded phases have ended, it is unlikely that the project will continue at the scale it used to.

As mentioned earlier, there is a lot of metadata. I introduce some types found in the data I am using and argue for considering some and not others for my analysis. First I discuss some of the situational and, afterwards, the personal metadata.

Audience size and gender is worth mentioning because a constant reminder of being filmed influences the informants' actions. There are two audience members in all three recordings, and in only one it consists of two women, in two others of mixed genders. In two cases the participants had a professional relationship with the audience members, and in one they had never met. The recordings begin right as the talking does, and ends just as participants decide they have talked for long enough, so there is no interaction between the

recorders and those recorded in my data. The medium is marked as private conversation in all cases even though the recorders are present but do not take part. In all three conversations the participants are marked as "Friend: members of the same group e.g. schoolmates". There clearly are differences in the participants' relationships, however, and they become evident in the conversations themselves in the form of topics of discussion. People tend to talk about things and interests they both share, such as a professor (classmates) or shared housing (roommates).

As for the location, in two cases the conversations were held in a lecturer's office and one conversation in a University seminar room. It is not known if the participants were familiar with the room and would thus associate it with a person whose it is, or if it is neutral territory (as much as a lecturer's room can be for a student). I do not consider the location in my analysis because the locations are not varied, and thus I cannot draw comparisons between them and speculate the difference it makes to the listeners' backchannelling behaviours.

Next I discuss shortly some of the personal data given. Places of birth and residence are of interest in one aspect only: the linguistic background. However, the metadata list languages the participants know, and if they speak, read, write, and understand them. The data do not specify how well the participants know these languages, so the amount of influence the non-first languages have on the linguistic competence spectrum is unknown. As the informants are all women, I need not long ponder gender differences.

> "THE EVIDENCE FROM ALL-WOMEN GROUPS IS THAT WOMEN VALUE HIGHLY THE ROLE OF LISTENING. THEY USE MANY MINIMAL RESPONSES, THEY DO NOT INTERRUPT IN THE SENSE OF PREVENTING A SPEAKER FROM FINISHING A TURN [...] AND THEY ACTIVELY ENCOURAGE OTHERS TO SPEAK."
>
> (COATES 1986, 154)

As Jennifer Coates argues, in search for backchannelling (or in her case *minimal responses*, see section 2.1.), the most fruitful informants would be women. As mentioned before, language is always changing and the language use of women has changed since most of the studies on minimal responses have been conducted, but trusting my own everyday experience and the influence of continuing pressure and stereotyping I assume the basic suggestion on women using more backchannels to hold water still. Thus, my material

consists of women-only interactions and therefore I will not be looking at gender differences or consider the effect gender might have on language use.

The recording years of the videos, two in 2005 and one in 2006, however, are not of interest for the purposes of this thesis since the time periods are so close to each other. When it comes to age, the metadata is uncharacteristically unspecific. Only the decade of birth is given, and thus strict distinction by age cannot be done. I do not believe this to be a problem since all listeners are born in the 1980s, and thus about the same age. Because of this and the small number of informants I cannot draw any conclusions about how age would affect backchannelling. The differences between individual listener actions are, most likely, due to individual differences of linguistic background and social relations.

3.2. METHODS

As stated in 2.1, I define *backchannels* as a verbal listener action that gives feedback and/or shows agreement, interest or involvement, and that is followed by the speaker continuing their turn, continuing the same topic. I do not follow a certain set of suggested backchannelling items, such as Coates' *minimal responses*, but do close reading on data chosen from SCOTS and infer the items functioning as backchannels there.

My method of collecting material comes from reading, listening, and watching the material. In reading the transcriptions that the SCOTS provides I am able to get familiar with the text and topics of discussion as well as get a preliminary idea of which backchannels seem to occur the most in the material. After familiarizing myself with the text I listen to the conversations, and here I am able to confirm whether the instances first thought to be backchannels, actually are. The intonation and pauses in audio aid in this. Lastly I move on to the videos from which I mark down gestures that co-occur with the predetermined and checked verbal backchannels. The conversations are situated in a room with participants sitting at a desk, so their legs cannot be seen. Only a part of the torsos and hands, and their heads are visible, so I will be only counting the gestures visible in these parts (such as a nod, leaning back and a wave of a hand).

In each conversation, I collect the relevant material into tables by person. This means five tables per conversation: two tables per listener (concerning first backchannels

and then gestures), and then one comparing the two listeners' backchannel use. In each table concerning a listener's backchannel use I list three most common backchannels, and if they occur in a longer section (i.e. a string), I include their position in it. A backchannel item under scrutiny could occur in the first position, meaning it starts the backchannel string, in the last position, where it is the final item, or somewhere in between, that is, in middle position (as distinguished also in Peters and Wong, 2015). Backchannels can also occur as standalones, and these could also be called *minimal responses*: they consist of only one word or item, such as *yeah* or *mm-hm*.

In collecting these backchannel instances, I also mark if they are a standalone or part of a string (and if, in what position), what reason do I have for calling them backchannels (their functions, such as continuer, acknowledgement, agreement and so on), and what gestures, if any, co-occur with these verbal backchannels. The data is put into an Excel table that is then fed into a code program called RStudio. The program is free, and allows easy rendering on information from tables into numbers: for example, a few lines of code can count how many times a listener uses *yeah* together with a nod. There is much more the program can be used for, but for the purposes of this paper, I only use it for its easy statistical summary functions.

In analysing the backchannel uses I compare the collected data tables, as shown in this thesis, with first their conversational partner and in Conclusions, against other listeners' tables. In these tables I either count instances or round up to the nearest 0.1%. I use percentages because they give the possibility of comparing listeners. For example, even when each listener has a unique amount of backchannel instances, the amount of standalones versus backchannel strings can be compared.

When there are multiples of same backchannel item in the same instance, I count these separately. For example, *yeah maybe, yeah probably, probably [laugh]* is one string with two instances of *yeah*, where one appears in the first and one in the medial position. As for the co-occurring gestures, I will indicate the most relevant one (visibly and meaning-wise), and apply that to the whole string. When it comes to counting gesture percentages per backchannel item, here I would count two instances of *yeah* as each co-occurring with a gesture, for example a nod, as they both are linked to the same gesture. This causes some weigh for some common gestures co-occurring with longer strings of backchannels, but I see this a better option than underscoring either the amount of backchannel items or commonness of certain gestures.

As what comes to the gestures themselves, there of course is more than one gesture at hand while in communication. However, for the sake of studying the most common ones, I focus on the most relevant gestures, as mentioned before in section 2.2.2, according to Goffman's *attentional tracks*. I also discuss some secondary gestures, such as nervous tics, as possible points of interest but do not count them into the tables and the final analysis. Seeing as my data rises from the conversations I study I do not rely on previous lists of gestures but look at the data themselves. This approach to gestures is the same as the approach towards verbal backchannels. This method takes time and is the reason I study only three conversations. This microanalysis demands more attention to small details and multiple rereading and re-watching of material, but it is this type of qualitative work that is needed to conduct a grass root type of study that gives insight into different phenomena. I do note, however, that because of the small amount of data I cannot expand my results and cannot draw conclusions about backchannel use in different or certain groups.

4. ANALYSIS

In the following subsections I analyse the conversations one by one. I begin by giving a short report on the general flow and impression of the conversation. Then I turn to the statistical information I have collected and discuss those by person. These, along with a descriptive report of the videos, hopefully paint a clear picture of the power of listener action over the conversation flow.

4.1. CONVERSATION 1

The informants in the first conversation are schoolmates who volunteered to participate in the filming of conversations for the SCOTS. They are around the same age, Listener 1 (hereon L1) born in Scotland and Listener 2 (hereon L2) in Germany. Neither of them knew the recorders beforehand. Since all given information about the conversationalist's

relationship is that they are schoolmates, we do not know their exact relationship status. They seem to be friends or acquaintances that have met a few times during common university classes since they talk about same assignments and courses.

The conversation starts off slowly and awkwardly, and both often stare at the table between them. Although L2 suggests few topics, they seem to be shot down by L1's unenthusiastic responses of delayed and quiet *mm*s and stares at the table. As Potter and Wetherell suggest, a delayed response often suggests a normatively dispreferred answer (1987, 84-5), which here would be an uninterested response – since the whole situation is built on the assumption that the two have a conversation. When looking closer, however, there are many small nods that do keep the participants talking, although with relatively long pauses in-between topics. However, these gesture-only backchannels are not what I am focusing on. There also seems to be another thing to consider: the effect of outsiders. The number of self-help advice on the Internet concerning the topic indicates that many people are nervous about speaking in public or in front of an audience. This, on top of being filmed, seems to make the participants more shy and that comes across as them being uninterested or distant. Nervousness affects people in different ways, and in this conversation it can clearly be seen: for example, L1 is quiet in her responses while L2 laughs a lot.

At around twelve minutes, after two of about six second pauses, the participants seem to find a common topic (books and movies), and the conversation starts to flow more smoothly. L2, although having laughed many times before due both to nervousness as well as amusement, becomes more animated in gestures towards the end of the conversation.

4.1.1. PROFILING LISTENER 1

Below I present two tables concerning Listener 1. Table 1.1 describes verbal backchannelling, including only the three most used ones. Here the backchannelling instances are divided between standalones, where backchannels occur only by themselves, and strings, where they are a part of a larger backchannel utterance. Following Peters and Wong (2015), I divide the strings into three positions: first, medial and final. First is what begins a backchannel utterance, final is what ends it, and everything else falls into the medial category. I count backchannels to belong in a string when they either comment on

the same speaker utterance or occur extremely close in time, and thus are to be a part of same thought and action - and thus the same instance of backchannelling.

Table 1.2 shows the most common gestures that co-occur with the aforementioned verbal backchannels. Here I present the three most common gestures and give percentages to show their commonness. I acknowledge that in some cases where the number of occurrences is small this seems unnecessary, but it allows comparisons to be drawn between different instances. I present these two tables together to facilitate comparison and to give a more complete picture of listeners' backchannelling actions.

TABLE 1. L1'S MOST COMMONLY USED BACKCHANNELS

	Standalone	**String**	**String position**			**All BC types**		
	Total	**Total**	First	Medial	Final	**Total**	Standalone	String
mm	14	0	0	0	0	**14**	100%	0%
yeah	5	0	1	0	0	**6**	83.3%	16.7%
oh	2	2	2	0	0	**4**	50%	50%

TABLE 2. L1'S MOST FREQUENTLY CO-OCCURRING GESTURES

BC	**Gesture 1**	**%**	**Gesture 2**	**%**	**Gesture 3**	**%**
mm	nod	28.6%	looks at table	28.6%	nongesture	21.4%
yeah	nod	33.3%	nongesture	33.3%	looks down	33%
oh	eyebrows rise	50%	eyebrows scrunch	25%	head forward	25%

L1's verbal backchannelling is nicely reflected in the table above. She is not really invested or interested in the topics suggested by L2, which can be seen in the few (relative to the 38 minutes of recording) and standalone (as opposed to complex string) responses. The longer the string, the more it highlights listener's interest. From the three most common backchannels it can clearly be seen that L1 prefers, in Coates' words, minimal responses. *Mm* is clearly the preferred verbal backchannel. Depending on the gesture co-occurrence, it can either indicate mainly a continuation or an (dis)agreement function. It occurs *only* as a standalone, which would suggest mainly continuer function: a more agreeing choice would be *mm, yeah*, for example. In example 2 below I use BC to refer for 'backchannel'.

EXAMPLE 1. L1'S USE OF *MM*

L1: I don't shop on a weekend.

L2: No, that's just, it's just too much.

L1: **Mm** [[nongesture / stares at the table]]

L2: I think.

EXAMPLE 2. L1'S USE OF *MM*

L2: Um Um, basically, good shopping opportunities. //Well, I don't have the money//

L1: //**Mm**// [[cocks head to left, BC long, disagreeing]]

L2: to [laugh] actually go shopping, but um, [tut], it's um, I think it's good that it's quite a big city, coming, like, coming from Hamburg and stuff. Um, where did you live before you went to Australia?

Yeah is mainly used as a standalone as well, except for *yeah, yep* where the verbal utterance itself seems to be agreeing, but tone indicates disinterest instead. In L1's backchannel use the meaning of *yeah* is strongly tied together to the tone it is said in, and the co-occurring gesture. Interesting things about this backchannel item is that although it seems to be the focus of quite many studies (as already mentioned before), it is used only six times by L1, and the more meaning-neutral *mm* is preferred, with a total of 14 instances.

Oh occurs only either as a standalone or in the frontal position. The item usually indicates surprise or interest, and as seen in table 1.2 above, all gestures co-occurring with *oh* do indicate these things. The rise of eyebrows and moving head toward the speaker show interest, and the scrunching of eyebrows, while possibly also disapproving, show (negative) affect toward the topic or speaker opinion, and thus interest and involvement. An example of *oh* indicating a positive reaction is down below.

EXAMPLE 3. L1'S SURPRISED *OH*

L2: Yeah, but she's she's trying to get in in a shape for her marriage.

L1: **Ah** [[slightly head upwards]]

L2: In August.

L1: **Oh!** [[(surprised, eyebrows rise, smiles]]

L2: So, sh- I think she's just trying to

As soon as we look at the most common gestures co-occurring with L1's most typical backchannels we begin to see exactly why it requires multiple viewings to recognize the listener's actions as shyness rather than outright disinterest. L1 nods quite a few times but when it comes to co-occurrences, there is not much movement there. For example only a bit over a fourth of gestures co-occurring with *mm* are nods, and the same amount is her turning to look down at the table. A fifth of co-occurrences are nongestural, which show no change of relevant, or meaning-making, positions. Nonrelevant positions here are things such as changing sitting position, which *can* indicate nervousness but usually does not add meaning to backchannelling itself. It can also be an indication of disinterest, distraction and/or boredom. Since the participants do not have their phones or other distracting items with them, distraction could be coming from L1's thoughts. Stress from dooming deadlines is a plausible explanation for students, and procrastinating and essays are even topics of discussion for both.

There is a very interesting backchannel-gesture combination in L2's use of *yeah* and nongesture. In quite a few studies *yeah* is shown to be one the most common verbal backchannel, so it is not a wonder that it shows here as well. It is, however, interesting how a word meaning *yes* can have such a passive and disinterested feel. After viewing the interaction multiple times it becomes clear that many of L1's backchannels are filling the silence. The conversation, for the most part, does not seem to spark enthusiastic, involved and interested continuers. It could be simply because the participants feel like there is nothing to talk about or because either one or both of them are camera shy and not willing to take the risk of starting a topic which would be disinteresting to either the other participant, or the future viewers of the film.

4.1.2. PROFILING LISTENER 2

As with analysing Listener 1, here too two tables are presented. The first table shows the three most used verbal backchannels and the second one the three most common gestures used with those three backchannels.

TABLE 3. L2'S MOST COMMONLY USED BACKCHANNELS

	Standalone	**String**	**String position**			**All BC types**		
	Total	Total	First	Medial	Final	Total	Standalone	String
[laugh]	43	10	3	0	7	53	81.1%	18.9%
yeah	11	26	14	7	5	37	29.7%	70.3%
mm	21	5	2	0	3	26	80.8%	19.2%

TABLE 4. L2'S MOST FREQUENTLY CO-OCCURRING GESTURES

BC	**Gesture 1**	**%**	**Gesture 2**	**%**	**Gesture 3**	**%**
[laugh]	rises head	22.6%	nongesture	17%	head away and back	15.1%
yeah	nod	56.8%	head tilt	13.5%	looks down	8.1%
mm	nod	61.5%	nongesture	7.7%	looks down	7.7%

L2 laughs a lot. She uses it mostly as a standalone backchannel, but does occasionally finish another backchannel with it. Since in every case it is accompanied by a smile, which by nature is social and empathetic, I do not count smile as a separate gesture. In western cultures laughter is often a sign of amusement or empathy, and at first it seems to be just this. After multiple viewings, however, it begins to seem like a sign of nervousness as well as a show of involvement and empathy. No matter what the intention, this backchannel together with the increasing animatedness of gestures towards the end of the conversation make L1 seem disinterested by comparison.

Yeah is mostly found in strings. Being a part of a backchannel string suggests that the other parts either add involvement by specifying what is agreed upon and/or is

associated with laughter, as seen for example in *yeah maybe, yeah probably, probably [laugh]*. Here the speaker (L1) has suggested, that "[m]aybe that's why they have a stop at Edinburgh, cause it might be one of the main //stations,//". L2 agrees with the proposition with *yeah*, specifies their agreement to *maybe*, and even highlights the possibility of this reason by *probably*. Agreement is added by L2 nodding and smiling. Giving L2's laughing any one particular meaning is difficult because she uses it so often and in many different situations. During the conversation L2's laughter changes between amused, empathetic, nervous, and a silence filler. Her laughter is not loud, and nowhere near malicious even in the example below.

EXAMPLE 4. L2'S USE OF *[LAUGH]*

L1: Yeah. Oh, I'm finding difficulties breathing out my nose.
L2: [**laugh**] [[rises head]]

Laugh is here always accompanied by a smile. In this paper I do not count a smile as a gesture when it occurs together with laughter, because it would take attention away from other, sometimes more telling, gestures. For example, first facing away from the speaker (sometimes scrunching the eyes closed) and then looking back at them conveys disbelief and thus emotional involvement with the speaker's utterance, although turning away can at first be seen as a distancing act in itself. Rising one's head, or throwing one's head back, shows either surprise or amusement, seeing as the action is, at least generally seen as, spontaneous.

L2 uses *mm* mostly as a standalone. Neutrality of the backchannel allows for its flexible use in many situations, such as surprise and sympathy, as seen in examples below.

EXAMPLE 5. L2'S USE OF *MM*

L1: See Townsville's got a reputation for cyclones.
L2: **Mm**! [[eyebrows rise, tilts head toward L1]]
L1: Yeah, so when it starts raining really really heavily, I remember, I was just like standing there in the

middle of ehm, I don't know really, [...]

EXAMPLE 6. L2'S USE OF *MM*

L1: Cause I'd rather get like lots of them, in a pack, than just one, cause the library sells them, but they sell them for like fifty pence or somethin like that,
L2: **Mm** [[tone sympathetic, frowns]]
L1: from a slot machine. You know where you get the pink one?

Nodding is the most common occurrence with *mm* so it is clear that the backchannel is mostly used to agree, ask for continuation, and indicate a slight degree of interest. What does not fit in the table above is that, along with nongesturing and looking down, it also has the same amount (7.7%) of rising one's head. So although *mm* can be used to agree (nod), show disinterest (nongesture and looking down), it can also be used to show interest (head rise). The co-occurring gesture (and tone of voice), then, can be manipulated to convey different meanings and intentions (such as interest) without much specific effort (for example using complex backchannel strings).

More than half of all *yeah*s and *mm*s are accompanied by nods – and by association, show agreement with the speaker. Tilting head with *yeah* can change meaning depending on the tone of voice. A rising tone indicates a question while drawn-out vowels together with head tilt and possible scrunching of eyes indicates suspicion or incredibility. Looking down creates distance mentally and, arguably, emotionally as well. So even though *yeah* has built-in meaning, it too can be changed by changing the tone and gesture accompanying it. This is one reason why when studying interactions it is important to look at gestures *together* with verbal backchannels.

4.1.3. COMPARING DISCUSSANTS

In this section I compare the backchannel use of L1 and L2. One topic of interest in this particular conversation is the differences in the amounts of backchannelling between the listeners, which can be seen in Table 1.5. In this section I also give more examples from the conversation that show the differences between the listeners.

TABLE 5. COMPARING L1'S AND L2'S BACKCHANNEL USE

	Standalone Total	**String** Total	**All BCs** Total	**% of** Standalone	**% of** String	**% of BCs in Discussion**
L1	36	15	51	70.6%	29.4%	**25.4%**
L2	98	52	150	65.3%	34.7%	**74.6%**

Both listeners use substantially more standalone backchannels than strings. It has to be noted as well that L1's verbal backchannelling is only a fourth of all backchannelling in this conversation. An interesting statistical titbit is that about the same amount (26.4% to be exact) of the conversation's total backchannels is L2 laughing. Again, laughing is used for varied purposes by L2, and as seen in the example below, it probably makes a big difference in the participants' communication.

EXAMPLE 7. L2'S USE OF *[LAUGHTER]*

L1: Hm [[long, looks down on hands]] That's the same with er Townsville and Cairns; they look so close, but they're like
L2: **Mm** [[looks down at hands]]
L1: really far apart. But then in Brisbane and Townsville, you get like that much, and that's like sixteen hours //car.//
L2: //**Mm**// //[**laugh**]// [[cocks head]]
L1: //[laugh]// A lot!
L2: [**laugh**] [[nongesture / keeps looking at other]]
L1: So, yeah.
L2: [**laugh**] [[looks down at hands]]
L1: I think two hours by plane though, two or three hours.

Laugh shows involvement, and this in turn keeps the speaker interested and involved as well: why would you not try to keep going if you are found interesting?

EXAMPLE 8. GESTURE SHOWS INVOLVEMENT

L2: What did she say?
L1: She said, eh, mm, she was talking about her family.
L2: **Mm** [[rises head, nods]]
L1: Eh, and then she says, after, sh- she only works till like one o'clock, or half-one [...]

In Example 8, L2 raises her head and is looking at the speaker, which would be taken as showing attention, and by extension, interest**.** In Example 1, however, L1 is staring at the table, hands still. These two instances of the same backchannel *mm* can, and are, interpreted differently by the current speaker, as can be seen from the utterances that come afterwards. In Example 8 the speaker continues the story, perhaps not having expected to be given a signal to continue (as could be interpreted from how she starts her turn with a hesitant *eh*). In Example 1 on the other hand the speaker hedges their previous statement ("I think"), which was most likely meant to give support to L1's statement in the first place.

Nongesture means that the listener is not moving, for example their head or hands. In this interaction nongestures are quite common, and even when coupled with such affirming backchannels as laughter or *yeah* (as is in L1's case), they, at most, indicate the basic function of speaker acknowledgement. Involvement and interest are low, feedback might be non-existent, and nongesturing might even express listener boredom – and thus a low desire for continuity. A casual interaction like this conversation would have ended very early on, but having agreed on being recorded, they clearly try to keep the conversation alive. At around 38 minutes, after a 4 second pause, they come to end the conversation.

EXAMPLE 9. END OF CONVERSATION 1

L1: So. Yeah.

[[PAUSE]]

L2: [laugh]

L1: Mm. Any more to talk about?

L2: Not really.

In the example above, note that L2's laugh is not a backchannel because it does not refer to the previous utterance but is a reaction to the long pause. The topic also changes, which means that the continuer function has not been filled.

4.2. CONVERSATION 2

Discussants in the second conversation are friends and flatmates. This fact is clear from their conversation topics: they talk about taking out the recycling, how the Listener 4's (L4) sister wants to move in for the summer if Listener 3 (L3) went to see her family, and they have common friends and a third roommate who has not shown up at her work. They are both born in Scotland and speak Scottish with their friends.

The conversation is easy-going, and discussants acknowledge from the get-go that they are being recorded. L4 has a pen in her hands and fiddles with it throughout the conversation, and L3 starts off with her arms crossed to gradually resting them on the table and finally leaning on it. She also fiddles with her sleeve quite a bit. L4 mainly looks down at her hands while backchannelling, whereas L3 almost exclusively looks at L4. The conversation flows effortlessly and without relatively long pauses, and turns are changed amicably and without either being left in the dust.

4.2.1. PROFILING LISTENER 3

In the tables below we see L3's three most common backchannels and the three most used gestures associated with those backchannels. As with L2, the most common backchannel is

laughter, but as I describe later on, it does not seem to be caused by nervousness but rather is a side effect of finding their conversation amusing.

TABLE 6. L3'S MOST COMMONLY USED BACKCHANNELS

	Standalone	**String**	**String position**			**All BC types**		
	Total	Total	First	Medial	Final	**Total**	standalone	string
[laugh]	39	16	4	0	12	**55**	70.9%	29.1%
yeah	4	13	7	3	3	**17**	23.5%	76.5%
oh	3	12	10	2	0	**15**	20%	80%

TABLE 7. L3'S MOST FREQUENTLY CO-OCCURRING GESTURES

BC	**Gesture 1**	**%**	**Gesture 2**	**%**	**Gesture 3**	**%**
[laugh]	shakes	60%	nongesture	20%	nod	3.6%
yeah	nongesture	64.7%	nod	11.8%	head tilt	11.8%
oh	nongesture	53.3%	nod	13.3%	turns away	13.3%

Most of L3's laughter is standalone, and it is clearly the most common verbal backchannel she uses. Judging from her almost permanent smile she is in a good mood and is entertained by the conversation. Unlike L2, her laughter does not seem nervous, and because she keeps looking at L4, her focus is on her conversational partner and the topics discussed. She also shakes a lot during laughter, which is a strong physical reaction towards what L4 is saying – and thus a deeper involvement and amusement. There are two interesting backchannels in this conversation, one including L3 laughing, as an example of interest and involvement below.

EXAMPLE 10. L3 SHOWING INVOLVEMENT

L4: […] and it's like ehm, you know, like the chivalric code that he's got to be courteous, and he's got to be eh, you know, polite and //he's got to be//
L3: //**[laugh]// //He's got to say please and thank you. [laugh]**//
L4: //[?]handsomely[/?] and all this.// //[laugh] You know?//
L3: //**"Ha-ha, I shall slaughter thee, please!"**//
L4: I should, you know, rise again, after I die, I just don't think, you know [laugh] it would be highly appropriate. […]

In the instance above, L3 is clearly involved in the conversation. She laughs, comments ("He's got to say please and thank you") without interrupting the speaker's story and even pretends to be the character in question by swinging an imaginary sword and giving the character imaginary lines ("Ha-ha, I shall slaughter thee, please!"). Before dismissing L2 *only* as a nervous and disinterested and L3 as an infinitely fascinated listener, however, it is important to mention that all people have different ways of expressing themselves: some people are by nature more expressive in their movements than others. The level of participation in Example 10 is not seen in the two other conversations under study, and can most likely be attested to both topic of conversation and close friendship where these types of interactions are normal.

Like the case with L2, over 70% of L3's *yeah*s occur in a string. As mentioned earlier, *yeah* in a string usually tells of agreement *and* interest or involvement, since the other parts of the string give focus to the exact parts of what is being agreed with. In the example below, L3 is agreeing with the claim that L4 would not be happy to spend her summer in darkness. *Ah*, for example, would have been a show of understanding or new but suspected information and *oh* of surprise. By letting on that she knows the claim to be a fact L3 shows that she knows L4 enough to know this information about her.

EXAMPLE 11. L3'S USE OF *YEAH* IN A STRING

L4: But I just, I don't know how happy I'd be, spending my summer in darkness. //You know, it's//
L3: //**Yeah, that's true**.// [[tilts head, looking forward]]

L4: at least in winter it's dark and cold, well, admittedly in summer it's dark and cold and miserable too, but I think I'd like to be there occasionally when the sun //shines.//

There are only two fewer *oh*s than *yeah*s used by L3. *Oh*, too, occurs mostly in strings, and, as can be seen from the table above, most often (83.3%) in the first position.

EXAMPLE 12. *OH* IN A STRING AND AS A STANDALONE

L3: //Are they a real// band, or are they just like a Uni //band?//
L4: //I don't know, I// think, well I think they're a real band, but I don't know how much of it's a band, and how much of it's a DJ. [sniff] //I think it might be//
L3: //**Oh right**.// [[nongesture, nods afterwards]]
L4: one of those fine-line DJ-band types. //[laugh]//
L3: //**Oh**// [[long, turns to look forward]]

What comes to secondary gestures that I do not take into account in the tables, L3 fiddles a lot when laughing. This fiddling cannot really be ascribed any meaning aside from perhaps nervousness from being recorded. Fiddling can also be a habit, or used to help focus on the conversation. Jessica Hullinger quotes Roland Rotz and Sarah D. Wright in her article *The Science of Why We Fidget at Work*: "If something we are engaged in is not interesting enough to sustain our focus, the additional sensory-motor input that is mildly stimulating, interesting, or entertaining allows our brains to become fully engaged and allows us to sustain focus on the primary activity in which we are participating." (www.fastcompany.com). Their book *Fidget To Focus: Outwit Your Boredom: Sensory Strategies For Living With ADHD* focuses on people with ADD (Attention Deficit Disorder, which is the preferred term to ADHD at this time), but even on their website Rotz and Wright state that they hope that the coping strategies outlines in their book will help everyone.

The quote above might unfairly be interpreted that if a listener is fiddling, they are bored. Long conversations, such as those that are analysed in this paper, demand effort and concentration from all participants. Fiddling helps listeners to keep the focus on the speaker, and some people naturally fiddle more than others.

4.2.2. PROFILING LISTENER 4

The tables below show the most common backchannels L4 uses in this conversation, as well as the most common co-occurring gestures. There are two main things to notice: one, that L4's most common backchannel is laughter, and two, that one backchannel to get into this list is *no* – the one word Bjørge wanted leave out of the definition.

TABLE 8. L4'S MOST COMMONLY USED BACKCHANNELS

	Standalone	**String**	**String position**			**All BC types**		
	Total	Total	First	Medial	Final	**Total**	standalone	string
[laugh]	19	9	3	0	6	**28**	67.9%	32.1%
no	1	5	2	2	1	**6**	16.7%	83.3%
uh-huh	5	0	0	0	0	**5**	100%	0%

TABLE 9. L4'S MOST FREQUENTLY CO-OCCURRING GESTURES

BC	**Gesture 1**	**%**	**Gesture 2**	**%**	**Gesture 3**	**%**
[laugh]	shakes	50%	head thrown back	21.4%	shakes head	10.7%
no	shakes head	33.3%	nongesture	33.3%	looks away	16.7%
uh-huh	nods	60%	looks at L3	40%	-	-

Laughter is clearly L4's most used backchannel, even though she used only half as much as L3 (28 versus 55 times). However, they use it very similarly: about 70% of laughter is standalone, and in those around 30% used in strings, it is twice as much used in the final than in the first position. An example of both instances is down below.

EXAMPLE 13. L4'S LAUGHTER

L3: I'm currently writing the last essay //of my university career. [laugh] I feel like I should end on a peak,//

L4: //**[inhale] Aw, no! So jealous! [laugh]**//

L3: but I don't think it's gonna //happen.//

L4: //**[laugh]**//

L3: [inhale] //It'll be a bit of an apology.//

Here *no* is also seen. As it is part of a backchannel string, I consider *no* to be a backchannel item. Even in text format this instance clearly indicates involvement and signals that the listener is jealous of the speaker's situation – as also highlighted by the rest of the string, "So jealous!". We have seen that *yeah* and nodding, both indicating agreement, usually co-occur – and in L4's case, the same is observed with *no* and shaking one's head. And when there is no gesture at all (as is the case in a third of all the instances of *no*), there is tone to indicate or add meaning, for example a lengthened vowel to show disbelief or added involvement.

EXAMPLE 14. UH-HUH AS A STANDALONE AND FILLING AS BACKHANNELLING

L3: It's eh, och, it's usually quite good. They don't get the best turnout . The music's really good, //so it's.//

L4: //**uh-huh**//

L3: I think I'll go along after the gig on //on Sunday night.//

L4: //**Sunday**.//

L3: If I've got any money left. //[laugh]//

L4: //[exhale]// Don't talk about money.

L3: I know.

In the example above there are three things I want to discuss. In the case of L4's use of *uh-huh*, all standalones co-occur with gestures pertaining to L4's head or gaze: a nod, or

turning to look at L3. Nodding implies agreement whereas turning to look at the other highlights that the listener is focused on the speaker – thus indicating interest. Another case of showing interest is finishing the speaker's utterance (as with "Sunday" above). In many cases filling the missing or delayed utterance is a helpful gesture and does not result in taking the turn. This is also the case here, where L3's hesitation (unfortunately not visible in the textual form) causes L4 to aid her by finishing the utterance. The last thing to note here is the second to last line: this could have been a backchannel if the speaker had continued her story, but in this case she very clearly responds to the now-speaker turn with "I know".

4.2.3. COMPARING DISCUSSANTS

In this section I compare L3 and L4's use of backchannels, their differing types, and backchannelling styles. Compared to the first conversation, the second one trots along more smoothly with both participants adding to topics and providing more complex feedback, that is, backchannel strings.

TABLE 10. COMPARING L3'S AND L4'S BACKCHANNEL USE

	Standalone	**String**	**All BCs**			**All BCs**
	Total	Total	Total	Standalone	String	**Total %**
L3	66	44	100	66%	44%	**65.8%**
L4	36	16	52	69.2%	30.8%	**34.2%**

Although L3 backchannels almost twice as much than L4, they both use about the same percentage of standalone and string backchannels. Also something to note is that compared to L1 and L2, L3 and L4 use more backchannel strings. As mentioned before, using strings indicates that the listener is more involved by giving more feedback and attention.

In the Example 10, L3 uses inventive backchannel strings, and they are not the only ones. In the extract below, in Example 15, she comments on L4's story about a TV programme. These long strings are treated as backchannels by L4, since she continues her story. Since these backchannel strings are long and inventive, some might stop their turn and listen what L3 has to say, but given how this situation is treated by both

conversationalists (notice how L3 does not try to take the turn by, for example, repeating a single backchannel multiple times) it seems as this type of exchange is usual and natural. It is exactly this feeling of familiarity between the two that gives the impression of close friendship.

EXAMPLE 15. L3'S INNOVATIVE BACKCHANNEL STRINGS

L4: So that was one channel, and they we had, eh, what did we have, like Home Improvements, but it was just, you know, some neddy housewife that was having her house re-done, and she was like "oh, it's gorgeous!", //[laugh] and all this, and it was horrible, and you know,//
L3: //**[laugh] Let's paint it bright pink! [laugh]**//
L4: ghastly and yuk, //and then eh//
L3: //**"Look we made this picture out of tin-foil"**.// //**[laugh]**//
L4: //and then they had eh// What's his name, that actor? […]

Looking at the backchannel items themselves one thing grabs my interest: while both L3 and L4 use laughter the most (and co-occur mostly with their body shaking showing amusement), L3's second most common item is *yeah* while L4's is *no*. *Yeah* co-occurring with nongesture indicates moderate interest and mostly continuer function (of course depending on other string items in backchannel strings), whereas *no* with shaking one's head (or nongesture) with a long vowel gives the feedback of involvement and interest. Together with L4's innovative backchannel strings, her use of backchannels shows that she is an involved listener, and that (in her opinion at least) L3 is a good storyteller and pleasant to listen to.

4.3. CONVERSATION 3

The third conversation feels like a two-way interview because the backchannelling, even at first watching, feels really "clean" and by the book. The listeners' relationship is listed as friends and members of the same group, but their conversation topics, such as listing what

they have studied before, indicates that they are more acquaintances than friends. There are three big giveaways in the way they communicate that reveals that they are not close. First, the conversation consists mainly of question-answer sequences, which is typical of interviews and first meetings. Second, most of the occurring backchannels are standalones, and the most common gesture across all backchannels is a nod. Standalones typically do not add any information or underline special interest, but more likely are used to encourage the speaker to continue. This could indicate that the listener does not have anything to add or contribute. Third, there is barely any talk over each other. In some cultures, for example in Finland, it is normal to wait until the previous speaker has finished their turn completely, but in English-speaking countries the continuous flow of conversation seems to be more prominent, and silence is considered rude and/or a sign of detachment. In interviews, however, it is important for the questioner to allow the speaker to finish their turn in order to gather as much information as possible, and this is why this conversation, in English, gives off the feeling of a bare-boned interview.

Listener 5 was born in Luxembourg and she speaks, reads, writes and understands English, French, German, Luxembourgish and Castilian, whereas L6 only lists English as her language. The differences in language background can explain the great amount of nodding – encouragement for the speaker or an assurance that the listener understands what the speaker says.

4.3.1. PROFILING LISTENER 5

Listener 5 (L5) knows multiple languages and cultures, as she was born in Luxembourg and speaks five languages. She has a varied linguistic and cultural background, and this gives her a large pool of backchannelling techniques to draw from – albeit it seems she does not utilise that when speaking English. Uncertainty in one's linguistic abilities can factor in a person using simple, common and basic backchannelling items, such as *yeah* and *uh-huh*, instead of complex backchannel strings that comment on the topic or the speaker's (or reveal the listener's) stance.

The table below shows L5's most common backchannels. As we can see, the most studied backchannels are there: *mmhm*, *uh-huh*, and *yeah*. The most frequent one, *mmhm*, appears in total 69 times whereas *uh-huh* only 27 times. Seeing as *yeah* is in the third place

with 19 occurrences, the jump from 27 to 69 seems like a lot. Looking at the gestures, however, the specific function of these backchannels does not seem to change.

TABLE 11. L5'S MOST COMMONLY USED BACKCHANNELS

	Standalone	**String**	**String position**			**All BC types**		
	Total	Total	First	Medial	Final	**Total**	standalone	string
mmhm	52	17	12	0	5	**69**	75.4%	24.6%
uh-huh	17	10	5	1	4	**27**	63%	37%
yeah	5	14	1	3	10	**19**	26.3%	73.7%

TABLE 12. L5'S MOST FREQUENTLY CO-OCCURRING GESTURES

BC	**Gesture 1**	**%**	**Gesture 2**	**%**	**Gesture 3**	**%**
mmhm	nod	78.3%	head up	8.7%	nongesture	8.7%
uh-huh	nod	70.4%	head up	11.1%	nongesture	11.1%
yeah	nod	63.2%	head up	15.8%	nongesture	10.5%

As is quite clear from the tables above, L5 clearly prefers nodding with the most common verbal backchannels she uses. An interesting thing to note is that she uses the same gestures with the three most common verbal backchannels – and about the same amount. Either these gestures are tied to how she backchannels in other languages, they are connected to the flow and feel of the conversation (an interview, lots of agreement), or they indicate her unfamiliarity with English spoken situations and that she takes refuge in these common and agreeable gestures.

Although not visible in the tables, an interesting fact is that L5 also laughs: in total 16 times, 13 of them are standalones. In three backchannel strings laughter is in the final position. When it comes to gestures, in 5 of the laughter cases her body shakes, in five she nods, and in 3 there are no gestures - so even here the pattern of mostly nodding appears.

EXAMPLE 16. L5'S BACKCHANNELLING

L6: So the- if they didn't have to do Maths they could potentially have done five //Highers just in//
L5: //**Highers, uh-huh**.// //**Yeah**.//
L6: //other things, so.// //I don't know if that's changing now or,//
L5: //**Mmhm**.// [[nods]]
L6: cause I think that was just, I know other schools didn't make you //do that.//
L5: //**Mmhm**.// //**[laugh]**//
L6: //It was just my school again! [laugh]// So, //yeah.//

EXAMPLE 17. L5 COMMENTS ON TOPIC

L6: computing and things like that. You know, when you get to fifth year now as well if you want to you can choose to do Psychology.
L5: **Uh-huh, that must be interesting**.
L6: And the school, there's a college quite near where I //live and//
L5: //**Mm**.// [[nongesture]]
L6: the school sort of sends you one afternoon a week to //college.//

In Example 16 there is one filling, “Highers, uh-huh”, where L5 helps finish L6’s utterance and at the same time conveys her understanding. There is also her most used backchannel, *mmhm*, used in its most common environment, as a standalone. Example 17 shows L5 commenting on L6’s speech, clearly indicating the wish for the other to continue by explicitly stating that the situation described is interesting.

4.3.2. PROFILING LISTENER 6

Listener 6 (L6), as mentioned earlier, only speaks English. An interesting titbit is that her parents were born in Glasgow, as well as she herself. So even though she has travelled

during holidays with her family, she has probably always lived in the same city which would limit her experiences and contact with other cultures and languages.

As can be seen in the table below, most of L6's backchannels are standalones. This fits in with the interview feel of the conversation.

TABLE 13. L6'S MOST COMMONLY USED BACKCHANNELS

	Standalone	**String**	**String position**			**All BC types**		
	Total	Total	First	Medial	Final	**Total**	standalone	string
mmhm	35	0	0	0	0	**35**	100%	0%
[laugh]	7	13	0	2	11	**20**	35%	65%
yeah	11	7	3	1	3	**18**	61.1%	38.9%

TABLE 14. L6'S MOST FREQUENTLY CO-OCCURRING GESTURES

BC	**Gesture 1**	**%**	**Gesture 2**	**%**	**Gesture 3**	**%**
mmhm	nod	74.3%	head up	14.3%	lean back	8.6%
[laugh]	nod	30%	shakes	30%	nongestrue	15%
yeah	nod	66.7%	head up	27.8%	lean back	5.6%

L6's gestures follow a similar pattern to L5: the most common gesture with *mmhm*, *[laugh]* and *yeah* is a nod or nodding. As seen in analyses concerning previous conversations, nodding is a sign of either agreement or an encouragement for the speaker to continue, and is fairly neutral in tone.

EXAMPLE 18.

F1148: so she kept just like screaming "I want my mum, I want my mum!" //They couldn't//
F1149: //**Oh no! And they couldn't understand her**.// //**[inhale]**//
F1148: //understand her. And then like we we- like we// when we like
realised she wasn't [inaudible], //we went down//

In Example 18 above, however, the seemingly neutral approach is clearly not the only thing going on. The exclamation "Oh no!" is a sign of involvement and interest in the story, as highlighted by the following filling of other's utterance. Inhaling as a backchannel is something I have not addressed before only because it is quite rare, but it clearly is a show of sympathy and involvement.

It is interesting, and rare, that a backchannel only be used in either as a standalone or in a string, but here L6 uses *mmhm*, her most common verbal backchannel, as a standalone. If I only looked at the instances of "mmhm" in the conversation transcription this would not be the case, but here microanalysis comes into play.

EXAMPLE 19. TAKING THE TURN

L5: The other two because, you know, she's just concentrating much more on //Math and Informatic you know, kind of thing.//
L6: //On Maths, mmhm.// How many years does she have left?
L5: Oh this is her //last year, uh-huh. [laugh]//

In Example 19 L6 finishes L5's utterance by "On Maths, mmhm", which could have been a backchannel if L5 continued her turn. Here, however, L6 takes the turn by asking a question – and L5 accepts this as a turn by giving an answer. L6's turn is *not* a backchannel even though it encompasses one of the most common backchannel items, because this utterance does not *function* as a backchannel.

4.3.3. COMPARING DISCUSSANTS

L5 and L6 use similar backchannels and more interestingly – their gestures are almost the same. They use mainly nodding and lifting their head (as a show of attention). This similarity could be accounted for by their unfamiliarity with each other. In these kinds of new situations speakers and listeners alike can rely on copying the others' gestures and speech

styles. If L6 knows of L5's diverse linguistic background (and that English is not her mother language) she can try to simplify her language use – and backchannelling – in order to make sure L5 understands her and not get confused by complex backchannelling.

TABLE 15. COMPARING L5'S AND L6'S BACKCHANNELL USE

	Standalone	**String**	**All BCs**			**All BCs**
	Total	Total	Total	Standalone	String	**Total %**
L5	71	43	114	62.3%	37.7%	**51.8%**
L6	68	38	106	64.2%	35.8%	**48.2%**

The table above shows that L5 and L6 backchannel almost an equal amount during this conversation. They seem to be on the same page about this conversation situation and eager to give space to each other to talk. If one tries to take the turn, the other one does not try to take it back. They ask questions back and forth, and even in the awkward situation of a long pause, if the other starts a topic it is not abandoned or ignored in any way, as can be seen in Example 18 below. Example 19 shows the beginning of the conversation, where the equality of the exchange is already established.

EXAMPLE 20. TOPIC CHANGE AFTER A PAUSE

L6: //Ah, alright, okay. [laugh]// Cause I remember seeing it.
L5: Uh-huh, yeah. [[Not BC because is followed by a change in topic]] [[Silence]]
L6: There's a n- a new erm place opened up over at Braehead, it's called Xscape or something. And it's a sort of
L5: Oh yeah, I've heard about that. //Uh-huh.//

EXAMPLE 21. BEGINNING OF CONVERSATION 3

L5: So, how old were you when you went to school?
L6: Ehm I was four when I started, ehm cause I started in August and I turned five in //October, so//

L5: //**Mmhm**.//

L6: really five, kind of thing. How about you?

The conversation ends abruptly, shown below, cut probably by the recorders probably for time purposes, since this conversation is already 32 minutes long. Until the end the exchange is like an interview, where one speaks and the other, by backchannelling, encourages them to continue.

EXAMPLE 20. END OF CONVERSATION

L6: //As you thought it was.// Like sometimes if you go somewhere and you you really like it, and you don't go back //for ages, you can//

L5: //**Mmhm**.//

L6: build up a

L5: **Yeah**.

L6: like build it up in your //head to be sort of better than it might have been.//

L5: //Mmhm.// [[Cannot be verified as BC]]

The last *mmhm* cannot be verified as a backchannel because we do not have the following speech or turn available. For an interested reader, I have put a longer example from near the end of this conversation in Appendix 2, which follows the basic structure of the discussion: the speaker tells a story while the listener backchannels, inspiring the speaker to keep talking.

5. DISCUSSION AND CONCLUSIONS

Involvement in a conversation and interest towards others and their topics can be seen in the listener's actions. For example, a keen listener usually shows interest and encourages

the speaker to continue their turn for a longer period of time, whereas a bored or uninterested listener can cause the turn or the whole conversation to wither (see section 1.1). A surprising finding for me was that even though one participant backchannels twice or even three times more than the other, it is not obvious from the transcription or the interaction until you count all the backchannelling instances. This is especially true in the second conversation, where the flow was even, both participated in talking and backchannelling, and neither seems to either hold back or take over.

TABLE 16. PERCENTAGE OF BACKCHANNELS USED AS STANDALONES PER LISTENER

BC	L1	L2	L3	L4	L5	L6
[laugh]	-	81.1%	70.9%	67.9%	-	35%
yeah	83.3%	29.7%	23.5%	-	26.3%	61.1%
mm	100%	80%	-	-	-	-
oh	50%	-	20%	-	-	-
uh-huh	-	-	-	100%	63%	-
no	-	-	-	16.7%	-	-
mmhm	-	-	-	-	75.4%	100%

Here I fall into the same trap as others – now I have a list of items as well, as shown in the table above. This list, however, is only applicable to the data at hand, and I do not recommend others to use it to find backchannels in other conversations. This list also includes laughter which I have not seen before, and *no* which is discounted by Bjørge in her definition of backchannel. Another interesting titbit is that out of six participants, *four* use laughter as one of the three most common backchannels. This is only beat by *yeah* with five. It could be argued that laughter falls more into the gestural category than verbal, but I decided to count it as a verbal instance for two reasons: one, it is verbal in the sense that in all these cases it is heard, and two, it has different gestural companions. For example, in the second conversation it mostly co-occurs with the body shaking, and in some cases there is no movement.

TABLE 17. PERCENTAGE OF BACKCHANNELS USED IN A STRING PER LISTENER

BC	L1	L2	L3	L4	L5	L6
[laugh]	-	18.9%	29.1%	32.1%	-	65%
yeah	16.7%	70.3%	76.5%	-	73.7%	38.9%
mm	0%	19.2%	-	-	-	-
oh	50%	-	80%	-	-	-
uh-huh	-	-	-	0%	37%	-
no	-	-	-		-	-
mmhm	-	-	-	-	24.6%	0%

What comes to backchannel strings, there are some clear differences in the placement of certain items. In the table above there is the percentage of backchannels used in a string, the positions of which can be found in the individuals' (L1, L2, L3, L4, L5, and L6) backchannelling tables. Laughter is mostly used in the final position in a string (where applicable in the data), whereas *yeah* can be in either the first or in the final position. This indicates that *yeah* can be used as a signal of backchannel string ending. This seems to be caused by its naturalness to the listeners, as it is one of the most used backchannels, and it is easily applicable for example where their thoughts drift off when speaker continues their story. *Oh*, on the other hand, is found almost always in the first position – although L3 makes an exception here, with two instances in the middle position.

When looking at gestures, the head's movements seems to have the central meaning-making part. The most obvious one is a nod, a gesture of agreement and/or understanding. Moving one's head also describes their attention: turning away or looking down creates distance by avoiding looking at the other, which is a sign of uncomfortableness, boredom or unvoiced disagreement. Throwing one's head back seems to be involuntary, or at least somewhat spontaneous: it often signals amusement and surprise, and an amused listener is more interested in the topic, which gives the speaker a good reason to continue their turn. Tilting one's head is a sign of confusion, and often interpreted as a request of either repetition or explanation.

Facial expressions are imperative with some backchannels. For example, in my data laughter is always accompanied by a smile. This is not counted in the gesture tables because there would not be space for any other gestures to peak through. Smiling seems to be a

natural part of laughter, and without it the backchannel is sarcastic which could work in certain contexts, e.g. when speaker talks about a co-worker who makes bad jokes. In those rare cases, however, it might be useful to consider that kind of backchannel a different type of laughter.

The videos of the conversations encompasses the upper bodies of the two discussants and a table between them. Not all their facial expressions are clearly visible, but some are. Rising or scrunching one's eyebrows, for example, is. These gestures depict surprise or disbelief, both of which indicate interest and involvement in the speaker's story.

5.1. ANSWERING RESEARCH QUESTIONS

The answer to this first question, patterns of co-occurrence, lies in the listeners' tables of backchannels and gestures. The only backchannel that is not used with a co-occurring nod is *no*, and *oh* only a few times. In all the other cases a nod or nodding can be found in some amount. In this section I discuss the questions of co-occurrence patterns and meaning enhancing, and later on address the question of listener profiling.

Laughter is always accompanied by a smile, as discussed above, and either the body shaking or head moving – either the head rises from looking down or is thrown back in amusement. As the third most common gesture during laughter L2 turns to look away, scrunches her eyes closed, and then turns back to look at L1. This gesture is unique to her, seeing as no other participant does this, and the commonness of it seems to suggest that it is a backchannel-gesture combination she uses outside this interaction as well.

Yeah co-occurs most often with a nod. With the built-in meaning, however, it can also be accompanied by nongesture and still be interpreted as agreement. *Mm*, on the other hand, is taken to be agreement only when co-occurring with a nod – and this is how it most often occurs. Looking away or down, thus creating distance between the discussants, even though still functioning as a continuer does not implicate agreement, and very little involvement or interest. *Mmhm* appears as one of the most used verbal backchannels only with L5 and L6, where, as discussed above, almost all of the gestures followed the same pattern: mostly nods, then head moving upwards, and third most common is nongesture. As with *mm*, nodding adds the meaning of agreement to the backchannel.

No is an interesting case of backchannelling. It is used for signalling sympathy and interest, and so the gestures co-occurring with it either strengthen the meaning (shaking one's head or turning away) or do not affect it (nongesture). In this case looking at the tone with which the backchannel item is uttered would be more lucrative, since for example a gasp indicates much more involvement than a monotone rendition.

Eyebrow movement, namely rising or scrunching them, is often linked with *oh* – either to indicate surprise or disbelief. In L3's case there is also nongesturing and nodding which most likely are linked to the other parts of backchannelling string that the *oh*s occur in. Nongesturing connected to *oh* seems odd since the word itself indicates surprise, and the gesture gives away nothing. However, as seen above with *no* and *mm*, not all gestures are there to strengthen the meaning of the word, but to work together with them.

Only L4 and L5 use *uh-huh* as one of their most common backchannels. L5 has an clear pattern of gesturing (nod, head up, nongesture) and L4 uses it only 5 times and only as a standalone. L4's most preferred gesture is a nod, and in the second place is turning to look at the speaker. What meaning these gestures bring to *uh-huh* is agreement and attention – as well as the invitation to continue.

The reason for looking at co-occurrences is, in effect, to highlight that they create different and/or new meanings. Verbal backchannels that have a relative neutral lexical meaning, such as *mm* or *uh-huh*, get their backchannelling function through intonation and gestures. This is why conversations through textual means, for example, lose some of their depth – because the listener as well as the speaker cannot recognize the exact meaning of the other's utterances. The gesture or rising one's head, eyebrows or chin and leaning back indicates surprise even when intonation does not. Intonation, another important part of verbal communication, is relatively easy to control unlike spontaneous gestures. This is why when studying lying, for example, special attention is focused on gestures and what they give away.

What comes to the third research question, that is, what backchannelling says about the listeners, I cannot make any tables or equations that predict their listener actions in future conversations. Although watching the interactions gives off a certain feel of what kind of listener these participants are, it must be remembered that the context affects the listening style. Different relationships have different dynamics and the presence of outsiders, not to mention the place of recording, all affect these interactions. Therefore, I want to highlight that the tables I have collected only apply to these specific conversations

and situations they have arisen in: in order to collect a more truthful or complete description of a person's listening style profile there would need to be many more conversations in different situations, in different environments, with different people and in different languages. What comes to the listening styles in these particular conversations, the analysis section of this thesis attempts to describe just that.

5.2. LIMITATIONS AND SUGGESTIONS FOR FUTURE STUDIES

As mentioned in the conclusions above, one methodological shortcoming for this type of qualitative study is, of course, the nongeneralizability of findings. Since there are just six informants of various linguistic backgrounds, we can only state that the results apply to the dataset used. Another problem of this study concerns subjectivity. It is a prevalent goal in many studies to strive to the ideals of 'hard sciences' and complete objectivity, but these cannot be wholly attained by one person only and not with the qualitative study methods I use. Therefore, I unsubscribe myself form these goals, and in this section focus only on what could be done to improve the type of study done in this paper.

As Sheida White states, "beginnings and ends of conversations, especially among strangers, are likely to be relatively problematic" (1989, 62). This can be caused by not having a shared interpersonal context: finding a topic of conversation both interlocutors are comfortable with can be hard work, as can be seen in the first conversation. A similar kind of problem, namely that of unnatural and hard work requiring conversation, can be caused by the interlocutors being aware that they are being filmed. Here an option would be to start the recording (or the analysis) sometime after the conversation has begun, so the discussants have time to establish a relationship and find a suitable topic.

Another cause for unnatural speech and backchannelling is the unnatural situation of *having* to strike up a conversation. Most interaction is, hopefully, voluntary, and the discussants are free to disengage when they wish. But in a similar fashion that Stanley Milgram theorized that people are likely to obey authority, it is likely that having promised to participate in a conversation the participants would not want to back down.

One methodological suggestion would be to play around with how to treat one backchannel item. Here I counted how many times an item occurred in the whole text – thus

when there were two *yeah*s in one backchannel string, I counted the two instances of *yeah* as two different occurrences. This has the advantage of realizing what backchannel items listeners use the most – but at the cost of multiplying gesture instances. For example, if a backchannel string includes two *yeah*s and one nod co-occurs, here I count two instances of *yeah,* both of which co-occur with a nod.

In the SCOTS transcriptions there are no pause lengths marked. This makes it difficult to transfer information from sound to text. Data about typical pause lengths in conversations can also provide useful so that misinterpretations can be avoided. One way to ensure all metadata wanted is present is to collect own material, but filming conversations takes time and monetary resources, not to mention a team to work on transcriptions. For the purposes of this paper I feel the data from SCOTS is enough, and more information on e.g. pause lengths only leads to more analysis and conclusions, which in turn demand more space and time to be properly addressed.

There is no one study that can fully explain the importance of listener action in a conversation. Tone, intonation, prosody, gestures, verbal backchannels, discussants' relationship – there are so many things going on at the same time that no one field of study can crack the code of how much listeners can affect interactions. Only by bringing multiple disciples together can we begin to comprehend the multimodality of listener action.

6. REFERENCES

ARIGA, Atsunori, & LLERAS, Alejandro. (2011). Brief and rare mental ‘‘breaks’’ keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition*. 118 (3), pp. 439-43. doi:10.1016/j.cognition.2010.12.007

BJØRGE, Anne K. (2009). Conflict or cooperation: The use of backchannelling in ELF negotiations. *English for Special Purposes*. 29, pp. 191-203.

BROWN, Pearson. (2007). Question Tags. Available at: https://www.englishgrammarsecrets.com/questionstags/menu.php [Accessed 19 October 2016]

COATES, Jennifer. (1986). *Women, Men and Language: A Sociolinguistics Account of Sex Differences in Language*. London & New York: Longman.

COATES, Jennifer. (2014). *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language.* 3rd ed. Routledge.

COATES, Jennifer. (2016). *Women, Men and Language: a Sociolinguistic Account of Gender Differences in Language*. 3rd ed. reissued. Oxon & New York: Routledge.

CRYSTAL, David. (2008). Two Thousand Million? *English Today*. 24 (1), pp. 3-6. doi:10.1017/S0266078408000023

EKMAN, Paul & FRIESEN, Wallace V. (1969). The Repertoire of Nonverbal Behavior: Categories, Origins, Usage and Coding. *Semiotica.* 1 (1). pp. 49-98.

Ethnologue. Summary by Language Size. Available at: http://www.ethnologue.com/statistics/size [Accessed 31 March 2016]

GUMPERZ, John J. (1982). The linguistic bases of communicative competence. In: TANNEN, Deborah. ed. *Analyzing Discourse: Text and Talk*. Washington, DC: George University Press.

HELLER, Monica. (2001). Discourse and Interaction. In: SCHIFFRIN, Deborah & TANNEN, Deborah & HAMILTON, Heidi E. *eds. The Handbook of Discourse Analysis.* pp.250-264. Oxford, UK: Blackwell Publishing Ltd.

HULLINGER, Jessica. (2015). *The Science of Why We Fidget at Work.* [ONLINE]. Available at: https://www.fastcompany.com/3044026/the-science-of-why-we-fidget-while-we-work [Accessed 15 December 2016]

JØRGENSEN, Marianne W. & PHILLIPS, Louise J. (2002). *Discourse analysis as theory and method*. London: SAGE Publications.

KENDON, Adam. (2004). *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.

KNIGHT, Dawn. (2011). *Multimodality and Active Listenership: A Corpus Approach.* London: Continuum International Publishing Group.

PETERS, Pam & WONG, Deanna. (2015). Turn Management and Backchannels. In: AIJMER, Karin & RÜHLEMANN, Christoph. eds. *Corpus Pragmatics: A Handbook.* Cambridge: Cambridge University Press. pp. 408-429.

PIPEK, Vojtěch. (2007). *On Backchannels in English Conversation*. Diploma Thesis, Masaryk University.

POTTER, Jonathan & WETHERELL, Margaret. (1987). *Discourse and Social Psychology: Beyond Attitudes and Behaviour.* London: SAGE Publications Ltd.

RÜHLEMANN, Christoph. (2016). *Backchannels in storytelling*. [Guest Lecture]. University of Helsinki, 3 February 2016.

SCHIFFRIN, Deborah. (2001). Discourse Markers: Language, Meaning, and Context. In: SCHIFFRIN, Deborah & TANNEN, Deborah & HAMILTON, Heidi E. *eds. The Handbook of Discourse Analysis.* pp.54-75. Oxford, UK: Blackwell Publishing Ltd.

SCHIFFRIN, Deborah & TANNEN, Deborah & HAMILTON, Heidi E. (2001). Introduction. In: SCHIFFRIN, Deborah & TANNEN, Deborah & HAMILTON, Heidi E. *eds. The Handbook of Discourse Analysis.* pp.1-10. Oxford, UK: Blackwell Publishing Ltd.

Scottish Corpus of Texts & Speech. Conversation 10: Two female students on travel and university. Available at: http://www.scottishcorpus.ac.uk/document/?documentid=800 [Accessed 01 April 2016]

Scottish Corpus of Texts & Speech. Conversation 11: Two female students chatting on various topics. Available at: http://www.scottishcorpus.ac.uk/document/?documentid=803 [Accessed 01 April 2016]

Scottish Corpus of Texts & Speech. Conversation 38: Two female students discussing school in the UK and Luxembourg and travel. Available at: http://www.scottishcorpus.ac.uk/document/?documentid=1575 [Accessed 01 April 2016]

SEIDLHOFER, Barbara. (2005). English as a Lingua Franca. *ELT Journal*. 59 (4), pp. 339-341. doi: 10.1093/elt/cci064

SHANNON, Claude E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*. 27, pp. 379-423.

STENSTRÖM, Anna-Brita. (1994). *An Introduction to Spoken Interaction*. London & New York: Longman.

TOTTIE, Gunnel. (1989). What does *uh-(h)uh* mean? American English vocalizations and the Swedish learner. *Instead of Flowers. Papers in Honour of Mats Rydén on the Occasion of his Sixtieth Birthday, August 27, 1989.* Stockholm: Almqvist & Wiksell International.

WHITE, Sheida. (1989). Backchannels across Cultures: A Study of Americans and Japanese. *Language in Society*. 18 (1). pp. 59-76.

WONG, Deanna & PETERS, Pam. (2007). A Study of Backchannels in Regional Varieties of English, using Corpus Mark-up as the Means of Identification. *International Journal of Corpus Linguistics.* 12 (4), pp. 479-509.

ZIMMERMAN, Don H. & WEST, Candace. (1975). Sex Roles, Interruptions and Silences in Conversation. In: THORNE, Barrie & HENLER, Nancy. eds. *Language and Sex: Difference and Dominance*. pp. 105-129. Stanford, CA: Stanford University Press.

7. APPENDIX

APPENDIX 1. MARK-UP CONVENTIONS USED IN THIS THESIS

SYMBOL	**MEANING**
[laugh]	laugh or giggle, adapted from SCOTS
// text //	overlapping speech, adapted from SCOTS
bolded text	backchannel
BC	shorthand for backchannel
[[text]]	gesture, or comment
[...]	part of text removed

APPENDIX 2. EXAMPLE OF INTERVIEW-LIKE CONVERSATION BETWEEN L5 AND L6

L5: //[laugh] Do you get to choose your subjects then for the Highers?//

L6: Yeah, y- well uh-huh. when you get to Standard Grades you have to choose eight subjects, //but really you only have//

L5: //**Mmhm**.//

L6: to choose five because it's compulsory to do Maths, English and either French or German, //whatever language you'd been doing.//

L5: //**Mmhm**.//

L6: You can choose the other five, so I did. And you had to pick at least one social science, //like either Geography,//

L5: //**Mmhm.// //Mmhm**.//

L6: //History, or Modern Studies.// So I did so I did Geography and History, Chemistry and Physics, and Art, //as well.//

L5: //**Mmhm**.//

L6: And then you can obviously, obviously you have to pick your Highers //as something you'd done in//

L5: //**Uh-huh**.//

L6: Standard Grades in. And in fifth year you have to do English and Maths still, so if you've, if you didn't get a good enough Standard Grade //to take a//

L5: //**Mmhm**.//

L6: a Higher in those subjects you have to take like an Intermediate //One or Two, which is//

L5: //**Mmhm**.//

L6: like a lesser qualification. But I don't know, when I was, just when I was doing my Highers, a lot of people were complaining about that, //because//

L5: //**Mmhm**.//

L6: there were a lot of people who really weren't so good at Maths, but because they were having to do it //in fifth year,//

L5: //**Mmhm**.//

L6: it meant they could only do four Highers […]